

学校编码: 10384

分类号_____密级_____

学号: 20091152767

UDC _____

厦门大学

硕 士 学 位 论 文

随机森林在图像分类和语义分割中的应用

Random Forests for Image Classification and Semantic
Segmentation

赵 自 明

指导教师姓名: 李翠华 教授

专 业 名 称: 计算机应用技术

论文提交日期: 2012 年 月

论文答辩时间: 2012 年 月

学位授予日期: 2012 年 月

答辩委员会主席: _____

评 阅 人: _____

2012 年 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1.经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ） 2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘要

随着网络的普及和多媒体技术的发展,目标识别技术已成为近年来需求和研究的热点。基于随机森林的目标识别方法在计算机视觉中是一类较新的方法,该类方法能够有效的处理大数据集,具有精确度高、鲁棒性强、易于实现、计算速度快等特点。然而,国内在该类方法上的研究仍较为空白,尚未出现总结性的文章。因此,研究基于随机森林的目标识别方法的工作,不仅对计算机视觉,模式识别等领域有一定的价值,而且希望能够对国内在该类方法的发展能够起到抛砖引玉的作用。

本文关注于目标识别任务中的图像分类和语义分割,主要研究基于随机森林的图像分类和语义分割的相关技术。主要的研究工作和创新点如下:

- 1、介绍随机森林理论相关知识,详细的给出训练算法和测试算法。
- 2、研究基于随机森林投票机制的图像分类模型,结合 MR8 滤波器组,将该模型应用于纹理图像的分类。在保持高分类精度的同时,较大的提高了分类的效率。
- 3、引入两种基于随机聚类森林的视觉词典。实现语义纹理元森林,联合基于金字塔匹配核的支持向量机,用于图像的分类。通过对语义纹理元森林和支持向量机的优化,进一步提高了分类精度。
- 4、在语义纹理元森林的基础上,设计基于类别概率分布信息的分割随机森林,串联语义纹理元森林和分割随机森林对图像进行语义分割。
- 5、根据随机森林易于并行化的特点,采用 OpenMP 多核编程技术并行化随机森林算法,较大的缩短了模型的训练时间和分类时间。

实验表明,基于随机森林的图像分类和语义分割算法在保持高分类准确度的前提下,能够快速的完成训练和测试。其中,基于 MR8 滤波器组和随机森林的纹理图像分类算法取得良好的分类效果。对随机森林算法的并行化优化,进一步提高了算法的效率。

关键词: 随机森林; 图像分类; 语义分割。

Abstract

With the development of internet and multimedia technology, recognition technology has become a research focus. Object recognition using random forest is a new kind of approach, which can handle very large scale dataset and has characteristics such as high precision, good robustness, easy implementation and fast computation. Yet domestic in this method of study on blank and the article summary has not appeared. For this reason, research work on object recognition using random forest not only has some value on computer vision and pattern recognition, but also in hopes of doing this to inspire domestic research on this method.

This paper pays close attention to image classification and semantic segmentation in object recognition task. The main research works and innovations are:

- 1 . Introduce random forest theory related knowledge, detail for training and testing algorithm.
- 2 . Study image classification model based on random forest voting mechanism. Combined with MR8 filter bank, apply this model to material image classification.
- 3 . Import two visual dictionary based on clustering random forest and introduce bag of word model. Implement semantic texton forest, this forest can be used to classify image followed with SVMs using pyramid matching kernel.
- 4 . Design region prior segmentation forest based on semantic random forest. Series semantic texton forest and segmentation forest to do semantic image classification.
- 5 . Parallelize random forest algorithm using Open MP multi-core programming language.

Experiment Indicates, image classification and semantic segmentations base on random forest can achieve high precision, also fast finish training and testing. Random forest and MR8 filter bank performs well in material image classification. Parallelize optimization random forest further improves the efficiency.

Keywords: random forest; image classification; semantic segmentation.

目 录

摘要.....	I
Abstract.....	II
第一章 绪论.....	1
1.1 研究背景和意义.....	1
1.2 相关研究工作.....	2
1.3 本文研究工作.....	4
1.4 文章组织结构.....	5
第二章 随机森林理论介绍.....	6
2.1 决策树.....	6
2.2 随机森林.....	8
2.3 极值随机森林.....	11
2.4 本章小结.....	13
第三章 基于随机森林投票机制的图像分类.....	14
3.1 随机窗口.....	15
3.1.1 提取.....	15
3.1.2 编码.....	16
3.2 模型介绍.....	16
3.2.1 训练算法.....	16
3.2.2 分类算法.....	17
3.3 纹理图像分类.....	19
3.3.1 MR8 滤波器组.....	20
3.3.2 分裂函数.....	23
3.4 实验.....	24
3.4.1 图像库介绍.....	24
3.4.2 实验结果与分析.....	27

3.5 本章小结.....	32
第四章 基于聚类森林的视觉词典.....	33
4.1 视觉词典介绍.....	33
4.2 词袋模型.....	36
4.2.1 基于叶子节点的词袋.....	36
4.4.2 语义纹理元森林.....	37
4.3 SVM.....	39
4.3.1 SVM 基本理论.....	39
4.3.2 多类分类问题.....	42
4.3.3 金字塔匹配核.....	43
4.4 实验.....	44
4.4.1 图像库介绍.....	45
4.4.2 实验流程.....	46
4.4.3 实验结果与分析.....	48
4.5 本章小结.....	52
第五章 语义分割.....	53
5.1 基于语义纹理元森林的语义分类.....	54
5.2 不变性学习.....	55
5.3 图像级分类约束.....	56
5.4 分割随机森林.....	56
5.5 实验.....	60
5.5.1 实验流程.....	60
5.5.2 实验结果与分析.....	62
5.6 本章小结.....	68
第六章 多核编程模型.....	69
6.1 OpenMP.....	69
6.1.1 OpenMP 简介.....	69
6.1.2 OpenMP 编程模型.....	69

6.2 实验.....	71
6.2.1 实验流程.....	72
6.2.2 实验结果与分析.....	72
第七章 总结与展望.....	74
参考文献.....	75
研究生期间发表论文及课题研究成果.....	80
致谢.....	81

Contents

Abstract.....	II
Chapter 1 Introduction.....	1
1.1 Research Backgrounds.....	1
1.2 Research status.....	2
1.3 Main Job and Contribution.....	4
1.4 Organization of this paper.....	5
Chapter 2 Random Forests Theory.....	6
2.1 Decision Tree.....	6
2.2 Random Forests.....	8
2.3 Extremely Random Trees.....	11
2.4 Summary.....	13
Chapter 3 Image classification based voting mechainsm.....	14
3.1 Random sub-window.....	15
3.1.1 Extraction.....	15
3.1.2 Coding.....	16
3.2 Model Introdution.....	16
3.2.1 Training algorithm.....	16
3.2.2 Classification algorithm.....	17
3.3 Material Image Classification.....	19
3.3.1 MR8 Filter Bank.....	20
3.3.2 Split Function.....	23
3.4 Experiments.....	24
3.4.1 Image Database Introduce.....	24
3.4.2 Experiment result and analysis.....	27
3.5 Summary.....	32
Chapter 4 Visual Dictionary based on Cluster Forests.....	33

4.1 Visual Dictionary Introduction.....	33
4.2 Bag of Word Model.....	36
4.2.1 Bag of Word Model based on Leaves.....	36
4.4.2 Semantic Texton Forests(STFs).....	37
4.3 Support Vector Machine.....	39
4.3.1 SVM Theory.....	39
4.3.2 Muti Class Categorition.....	42
4.3.3 Pymarid Matching Kernel.....	43
4.4 Experiments.....	44
4.4.1 Image Database Introduce.....	45
4.4.2 Procedure.....	46
4.4.3 Experiment result and analysis.....	48
4.5 Summary.....	52
Chapter 5 Semantic Segmentation.....	53
5.1 Semantic Segmentation based on STFs.....	54
5.2 Learning Invariance.....	55
5.3 Image Level Prior.....	56
5.4 Segemntation Forests.....	56
5.5 Experiments.....	60
5.5.1 Procedure.....	60
5.5.2 Experiment result and analysis.....	62
5.6 Summary.....	68
Chaption 6 Muti-Kernel Programing Model.....	69
6.1 OpenMP.....	69
6.1.1 OpenMP Introduction.....	69
6.1.2 OpenMP Programing Model.....	69
6.2 Experiments.....	71
6.2.1 Procedure.....	72
6.2.2 Experiment result and analysis.....	72

Chaption 7 Concolutions and Future work.....	74
Reference.....	75
Publised Papers and Subject Research Achievements.....	80
Acknowledgement.....	81

厦门大学博士论文摘要库

第一章 绪论

1.1 研究背景和意义

在计算机视觉领域，目标识别是一项重要且富于挑战性的工作。目标识别包含了三方面的内容：图像分类，目标检测，语义分割。图像分类的目标是将一组包含相同类别目标的图像分组为同一集合。一张图像可能包含多个不同类别的目标，也可以不包含任何目标而仅以背景形式存在。目标检测既需要完成图像分类的任务，又要求标定目标在图像中的位置与尺度。语义分割将图像分离为连续的区域，同时为每个区域指定所归属的类别^[1]。经过科学家和工程师的不懈努力，这些技术在图像检索，医药，机器人，安全保障，交通，图像编辑等领域已有实际的应用。

基于计算机的目标识别系统因为缺乏人类关于视觉语义的高级知识而面临着严峻的挑战。“横看成岭侧成峰，远近高低各不同”，“毕竟西湖六月中，风光不与四时同”，形象地描述了目标在不同的视角，光照，尺度，姿态，部分遮挡等因素下所具有的各种形态。这些形态变化多端，使得描述相同类别目标的内部共性与不同类别目标的区分性变得极为困难。理论上探寻一个能够准确描述所有形态的计算模型可谓大海捞针。在实际应用中，大多数的识别方法基于有标注的图像库，提高图像训练库的质量和数量的方法可提升识别的准确度。一般而言，属于同一类别的训练数据愈多，模型就能愈完整的捕获该类别的内部多样性。对图像标注的信息量越多，则在训练阶段就越能准确的确定图像中的特征所归属的类别^[2]。不幸的是，对图像库的丰富信息的丰富需要大量的人力支持，所以，更多的学者关注于寻找有效表述目标特征的方法以提高识别的准确度^[3, 4, 5, 6]。

高等级的目标特征表示往往能带来识别准确率的提升，然而，高等级目标特征的提取一般具有一定的计算复杂度，如 SIFT^[3], Filter-Bank^[7]。对于海量的图像数据库，提取这样的特征需要较大的时间代价。在提取特征和量化特征之后，多种分类器可用于实现对特征的分类。在海量数据库中，分类器（如 K 近邻，神

经网络, SVM, Boosting) 的训练同样需要消耗大量的时间。

针对以上提出的几个难题, 视觉领域衍生发展出许多新的思想和方法。如何在保持识别准确度的前提下, 提高系统训练和识别的效率, 已逐渐成为目标识别领域发展的主要方向。

随机森林是一种高效且具有高区分度的分类器, 它在获得较高的分类精度的同时, 具有训练时间短, 分类速度快等特点。文献[8, 9]对随机森林的介绍使其在机器学习领域引发广泛的研究热潮。在Maree^[10], Lepetit^[11]等人的努力下, 随机森林被引入计算机视觉领域并诱导出了一大批基于随机森林方法的文献: [12, 13, 14, 15, 16, 17]。

由此可见, 基于随机森林方法的目标识别技术已在国外掀起一股“研究热”, 而国内随机森林理论的相关论述则寥寥无几, 尤其对随机森林在目标识别中的应用研究更是空白。因此, 本文对随机森林方法在图像分类和语义分割上的已有方法进行总结和改进, 是一个有意义的研究课题, 希望该研究能够对国内在该类方法上的发展能够起到抛砖引玉的作用。

1.2 相关研究工作

在图像分类领域, 文献[18]所提出的框架已被广为采用, 该方法首先提取感兴趣的特征点, 利用特征点周边信息将特征点量化为数值向量, 这些数值向量被用于模型的学习或分类测试图像。文献[10, 19, 20]采用相似的框架。文献[19]提出了基于决策树和随机窗口的图像分类方法。在训练阶段, 从训练库图像中随机地采样出固定大小的窗口作为特征, 基于简单的颜色空间量化方式将这些特征量化为数值向量, 这样的向量集合被用于训练随机森林分类器。在测试阶段, 从一张待分类图像中提取一个向量集合, 利用随机森林的叶子对这些向量分别进行投票, 根据投票结果选举产生图像最可能的归属类别。文献[10]在文献[19]基础上改进了随机窗口的提取方式, 使随机窗口的大小自由变化, 而在特征量化时再将窗口归一化为固定大小。这一改进提高了分类的准确度。文献[20]基于以上的框架, 对各种树族类方法的分类效果进行比较分析, 这些树族类方法包括单决策树, 随机森林, 极值随机森林, Bagging, Boosting。

词袋作为一种有效的信息表征工具,最先来源于文本分类领域^[21]。随后, Siciv 等人在文献[22]中首次将其引入计算机视觉领域,文献最大的贡献在于广义化了文本中的单词概念,使计算机视觉领域出现了新的名词:视觉单词和纹理元。

纹理元^[7, 23]和视觉单词。在图像分类和分割中具有强大的特征表现能力^[24, 25, 26, 27]。纹理元和视觉单词在本质上具有相同的意义。Filter-bank (高斯微分, 小波等)或不变性描述子 (SIFT等)在训练图像库中以稀疏^[14]或稠密^[16]的方式进行计算产生描述子向量,这些描述子聚类产生一个由视觉单词构成的视觉词典。K-Means是一种简单有效的聚类方法^[28],每个描述子采用最近临赋值算法选择与其最接近的视觉单词。这三个步骤耗时较多,占据整个算法大部分的运行时间,即使采用K-D树和分层聚类等方式进行优化也不能明显改善计算效率。

Moosmann提出一种更为有效的聚类方式:随机聚类森林^[14]。随机聚类森林上以简单的对描述子中的某一维与随机阈值进行比较作为节点的分裂属性,训练样本集在树上递归地分离为两部分,整个过程不存在大量耗时的环节。随机聚类森林中每棵随机树的叶子节点均被视为视觉单词。每一张图像选取一定数量的描述子,以深度递归的方式为每个描述子寻找其在每一棵树上所对应的视觉单词,从而形成一个视觉单词统计向量,训练库图像得到的向量集合用于分类器SVM训练,而测试图像则将向量输入分类器SVM, SVM的输出结果为该测试图像的归属类别。

Jamie在Moosmann的工作上进行扩展^[16],提出语义纹理元森林 (Semantic texton forest),它可以用于图像分类和图像的语义分割。节点分裂属性采用像素值比较的方式,避开采用计算量较大的图像描述子。语义纹理元森林既可隐式的分层聚类为语义纹理元 (除根节点之外的所有节点都可视为视觉单词),同时,通过叶子节点亦可获得显式的局部类别概率估计。在图像分类时,图像以滑动窗口的形式采集样本集,将样本集导入语义纹理元森林,得到一个树层次结构的词袋统计,分类器SVM对该词袋分类得到图像的归属类别, SVM采用基于层次结构的金字塔匹配核^[29]。

在图像的语义分割中,文献[16]引入分割随机森林,图像通过语义纹理元森林可获得纹理元的词袋统计图和类别概率分布信息图,分割随机森林的节点以两

种图的局部统计值作为分裂属性，生长完成的分割随机森林中所有树的叶子节点包含类别概率统计。一张图像可通过两个串联式的随机森林进行语义分割。

1.3 本文研究工作

本文关注目标识别问题中的图像分类和语义分割。主要做以下方面的工作：

1、介绍随机森林相关理论基础，阐述决策树，随机森林和极值随机森林等算法框架，比较训练算法的异同，并讨论其鲁棒性，分类精度和计算效率。

2、研究基于随机森林投票机制的图像分类模型，对随机森林的训练方式进行改进优化，实现该模型并在 COIL-100^[30]，ZuBuD^[31] 图像库上评估关键参数的选择对分类准确度的影响。

3、实现 MR8 滤波器组^[7]，将上述模型应用于纹理图像的分类。我们在 Columbia-Utrecht (CURET)^[32] 图像库上测试算法精度和计算效率。实验表明，该方法具有较快的训练时间和测试时间，同时保持较好的分类准确度。

4、探讨两种基于随机森林的词袋模型，其中一个模型以随机森林的叶子节点作为视觉单词，而另一模型则以随机森林的除根节点意外的所有节点作为视觉单词（纹理元）。我们实现后一种模型，并在 MSRC21^[25] 图像库上进行评估，通过优化实验中的关键参数，引入加权的平衡训练，我们取得比[16]更好的分类效果。

5、介绍语义纹理元森林和分割随机森林在图像的语义分割中的应用，引入图像级类别概率约束增强语义分割精度。我们实现语义纹理元森林和基于类别概率分布信息的分割随机森林，实验表明，分割随机森林相比语义纹理元森林具有更高的语义分割精度，通过参数的优化选择，我们获得了比文献[16]更高的语义分割精度。

6、采用 OpenMP^[33] 多核编程技术，充分利用随机森林中树与树之间的独立性，将随机森林并行化，进一步随机森林的训练和分类的效率。通过并行优化，我们实现了快速的图像分类和接近实时的语义分割。

1.4 文章组织结构

本文的章节安排如下：

第一章 绪论

主要介绍目标识别的发展现状和遇到的难题，说明基于随机森林的目标识别方法的国内外现状。追述了国外在该类方法上的相关工作，最后说明了本课题所做的工作与文章结构安排。

第二章 随机森林理论介绍

介绍决策树，随机森林，极值随机森林，详细的给出训练算法和测试算法。评价随机森林的显著性优势。

第三章 基于随机森林投票机制的图像分类

本章从随机窗口提取技术讲起，随后介绍基于随机森林投票机制的分类模型并将该模型引入纹理图像的分类中，最后在 COIL-100，ZuBuD 和 Columbia-Utrecht 等图像库上对实验结果进行比对分析。

第四章 基于聚类森林的视觉词典

本章给出基于随机聚类森林的视觉词典，对词袋模型的工作流程以及词袋模型中的关键分类器 SVM 进行了研究，结合随机聚类森林阐述金字塔匹配核的计算过程。最后在 MSRC21 图像库上进行测试与分析。

第五章 语义分割

本章给出基于随机森林的语义分割方法，串联语义纹理元森林和分割随机森林对图像进行语义分割。我们在 MSRC21 图像库上进行测试并对实验结果进行分析和讨论。

第六章 多核编程模型

介绍 OpenMP 多核编程模型，并对随机森林并行化之后的效率提升进行评估。

第七章 总结与展望

对本文的工作进行总结，并讨论本课题进一步的研究方向。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库